

Insights from developing and evaluating the NHS blood choices transfusion app to support junior and middle-grade doctor decision making against guidelines

Amanpreet Singh Dhesi¹ | Jeremy C. Wyatt² | Lise J. Estcourt^{1,3}  | Wendy McSporran⁴  | Shubha Allard⁵

¹Department of Blood, NHS Blood and Transplant, London, UK

²Wessex Institute of Health Research, University of Southampton, Southampton, UK

³Radcliffe Department of Medicine, University of Oxford, Oxford, UK

⁴Department of Blood Transfusion, Royal Marsden NHS Foundation Trust, London, UK

⁵Department of Patient Blood Management, NHS Blood and Transplant, London, UK

Correspondence

Amanpreet Singh Dhesi, NHS Blood and Transplant, NHSBT 75 Cranmer Terrace, London SW17 0RB UK.
Email: amanpreet.dhesi@gmail.com

Funding information

NHS Blood and Transplant

Abstract

Objectives: To: 1. Develop a CE-marked smartphone App to support doctors' concordance with transfusion guidelines in non-bleeding adult patients, emphasising informed consent and anaemia management. 2. Test App accuracy and potential to improve user decisions.

Background: Studies have shown inappropriate use of blood components and that most junior doctors own smartphones with medical apps.

Methods: A multidisciplinary team developed App screens and logic through an iterative process based on national guidelines. Thirty medical or surgical transfusion scenarios were developed based on national guidelines and each sent to Consultant Haematologist experts in Transfusion Medicine. To obtain a clinical consensus and exclude ambiguous scenarios, their independent decisions and associated certainty were compared. The consensus clinical decision was then compared with guidance from the App. To explore potential App impact on simulated user decisions, 26 junior doctors responded to five transfusion scenarios before and after access to the App.

Results: The Blood Choices App agreed with 91% (95% CI: 72%–99%) of expert decisions with a sensitivity of 100% (69% to 100%) and specificity of 85% (55%–98%). Excluding one malfunction scenario, the App had the potential to increase correct decisions by junior doctors from 83% (73%–90%) pre-App use to 96% (88%–99%) post (p -value 0.013), with 90% (67%–99%) saying they would use it in practice.

Conclusions: Transfusion guidelines can be converted into an App with potential to improve guideline concordance. However, evaluating such Apps is essential to understand their limitations, detect malfunctions and prevent harm.

KEYWORDS

decision accuracy, evaluation study, guideline implementation, simulated impact study, smartphone app

1. BACKGROUND

Clinical guidelines are developed to improve patient care by supporting healthcare professionals to apply the latest evidence-based recommendations and make effective use of resources.¹

Multiple audits by the National Comparative Audit of Blood Transfusion (NCABT) have shown large variations in transfusion practice and a lack of concordance with guidelines. A cardiac surgery patient audit showed considerable variation in the percentage receiving a red blood cell (RBC) transfusion ranging from 22% to 66%.² High

levels of RBC transfusion were also reported as inappropriate in the 2011 Medical Use of Blood Audit, with 20% of RBC transfused to patients with potentially reversible anaemia.³ This highlights the ubiquitous challenge of ensuring transfusion guidelines are implemented.

To reduce unwarranted variation and improve evidence-based transfusion practice, clinical guidelines have been published by the National Institute for Health and Care Excellence (NICE) and the British Society of Haematology (BSH). Past interventions to increase adherence with these guidelines have included: reminders, education, checklists and algorithms. However, expectations that these ongoing efforts would substantially improve adherence have not been met, as subsequent audits have only shown small increases in adherence.⁴

An innovative solution to improve guideline adherence and address inappropriate RBC use by doctors was therefore sought which takes advantage of the increasing medical use of smartphones.

1.1. Healthcare Smartphone App Usage

The use of Apps is professionally and socially acceptable,⁵ with up to 99% healthcare professionals owning and using smartphones during clinical practice.⁶ As healthcare takes place in multiple clinical locations, Apps can improve patient care where remote access to inform decision-making is required.⁵ Doctors endorse the development of Apps to support clinical practice and their education.⁷

A survey by Charani et al.⁸ reported 81% of junior doctors stated an App helped them adhere to local hospital policy. Panesar et al.⁹ transformed antimicrobial prescribing guidelines into an App concluding it is an effective and acceptable format to deliver guidelines. The App was utilised more than the pocket-guide previously used to promote guideline implementation. Increased usage was associated with the App being useful and easy to navigate. They also concluded that the App allowed users to challenge incorrect prescribing and 'break well-established behaviours'.⁹ This requirement for behaviour change is consistent within blood transfusion, where past audits of practice have shown a lack of guideline adherence and past interventions have also included development of paper aide memoires.

NHSBT is responsible for providing transfusion training to many healthcare professionals with wide use of case studies that can be used for simulation scenarios. Song et al.¹⁰ and Low et al.¹¹ reported evidence that during simulations, Apps had a positive effect on guideline adherence. However, the true benefit of evidence-based healthcare must be reflected within the clinical setting. As guidelines can be adapted into App format for easy access at the point of care, this is potentially a useful solution to improve clinical adherence with blood transfusion guidelines.

The Royal College of Physicians¹² guidance for Healthcare Professionals, stated that professionals must only use CE-marked Apps to support clinical decisions and check that Apps have been tested for accuracy and benefits in clinical use.¹² The publication noted that healthcare professionals are left open to increased risk if they use Apps without applying these safety criteria.

The aim of this pilot project was twofold. First, to explore how a free-to-download iOS and Android CE-marked App containing accurate RBC guidelines for non-bleeding adult patients, emphasising informed patient consent and anaemia management, could be developed. Second, to evaluate the potential of this app to improve doctors' decision-making.

The clinical content of the App focused solely on evidence-based RBC transfusion guidelines for stable, non-bleeding adult patients. The app was not intended to be used in the emergency setting.

2. METHODS

During the planning of this project there was a lack of established practice defining the best way to develop, implement and evaluate Apps in healthcare. Therefore, a logical and systematic approach to App design and evaluation was carried out. Since then, Wyatt¹³ published details on evaluating and improving the quality of clinical Apps, which this project was compliant with.

To facilitate doctors' ease of access to guidelines, the first objective was to transform lengthy guidelines into an App.

2.1. Creation of the Blood Choices App

The relevant guideline content was identified and transformed into flow charts that define the user journey based on patient data inputs. Any variable entered that indicated the patient fell outside the App's scope (e.g. patient age < 16 years old or massive haemorrhage) led to an exclusion message screen.

To check if the input variables led the user to the correct outcome screen, combinations of variables were developed into simple clinical case studies. Clinicians then confirmed if the outcome screen advice matched guideline recommendation for the clinical case presented. The user journeys, inputs and output screens were developed via paper-based hierarchy maps, web-based simulations and on smartphones.

An App working group was formed of, and engaged with, multi-disciplinary clinical teams including Transfusion Practitioners, doctors of different clinical grades and specialties, digital healthcare professionals and biomedical scientists. Junior doctors and experts in other relevant fields were also involved in the development through open and direct engagement by the working group, either on a one-to-one basis, through workshops or during breaks in their clinical induction.

To test the App's intended functionality, it underwent validation against a detailed, signed off (both clinical and technical) validation script which indicated every possible user journey, every App function and even functions which should not occur (e.g. entering text instead of a number). All requirements for development, validation and post market surveillance required by the Medicines and Healthcare Regulatory Agency were met. The App was then self-declared as being a Type-I medical device so could display the CE mark.

2.3. Testing App Safety and Accuracy

To test App safety, the first study evaluated the accuracy of advice given by the App. Between November 2016 and January 2017, 30 scenarios (Appendix S1) were created based on typical cases that were reported as receiving appropriate and inappropriate transfusions from past clinical audits and reports. Using the NHSBT Patient Blood Medical Team network, consultant haematologists with expertise in transfusion medicine were invited to volunteer to provide answers to these scenarios. There were no inclusion/exclusion criteria, and the 24 selected haematologists were those who volunteered their time. Independent external consultants were used to reduce any bias in scenario creation or App interpretations originating from the development group. To eliminate ambiguous scenarios, consultants were asked for their decision about each scenario (transfuse/do not transfuse) and for their decision certainty using a visual analogue percentage scale.

They were each asked to provide a decision for up to five scenarios with the aim to obtain a clear majority decision per scenario about whether to transfuse RBC or not. To achieve consensus, we required the overall mean certainty score to be above 50% (as certainty below this suggests an ambiguous scenario) or a greater than two-thirds majority consensus decision.

The consultants then used the App and recorded its output. The consensus clinical decision for each scenario was then compared to the advice obtained by the consultant using the App. Those scenarios with no App response recorded by the consultant or without a clear majority consensus were excluded to leave 23 scenarios.

In addition, those scenarios that had the lowest certainties or the most disagreement were excluded from the next study.

Sensitivity and specificity are used to assess the performance of investigations in a clinical setting. In this study, sensitivity measured the ability of the App to correctly conclude that a transfusion is required, and specificity was the ability of the App to correctly advise that a transfusion is not needed.

2.4. Study to test App impact on simulated decision making by junior and middle-grade doctors

The aim of this study was to explore the effect of the App on guideline adherence, using scenarios to simulate clinical use. The criterion for selecting the five clinical scenarios (S1, S8, S9, M15 and S7) for this study was that each scenario showed strong expert consultant consensus on correct management, rather than individual clinical significance or frequency of errors identified in previous audits.

In May 2017, an online survey (*SnapSurvey*®) was created to capture junior doctor responses to the five scenarios based on their current ability, knowledge and resources available. They were then directed to download the App and to respond to the same five scenarios, this time using the App. Demographic and App usage information was also obtained.

Working group members recruited volunteers (foundation year doctors and those who were not working in haematology/transfusion)

in their local hospitals through posters, requests at training sessions and one-to-one discussions. A prize draw was offered as an incentive to complete the survey. Response rate was not measured due to this varied approach to recruit this hard-to-engage group.

A chi-squared test was carried out on the overall results—including both junior and middle grade doctor decisions—to compare the number of correct verdicts with and without the App. The dataset fulfilled the assumptions for the chi-squared test. The test was conducted using Microsoft® Excel® for Office 365 MSO (version 16.0.11328.20492) 32-bit.

2.5. Information Governance

The App was designed not to collect any patient-identifiable data and only basic demographic details for doctors completing the evaluation survey were collected, without any identifiers.

Google analytics was built into the App but was not used in these evaluations.

3. RESULTS

The Blood Choices App was built for Apple and Android Smartphones and was CE-marked and free to download. The App was not publicly launched and was only promoted to users involved in this study on their personal phones. They were asked to delete it once they had used it for this evaluation. Every time any changes were made to the App it was taken off the market and relaunched for pilot work only. The results discussed in the following sections indicated that further changes and testing were required, but as resources were limited to pilot phase only, the App was not placed back on the market. Screen shots that demonstrate the App's functionality can be seen in Appendix S1.

As part of the validation process for CE marking, the App was tested on different smartphone platforms and with different operating systems to ensure it gave consistent advice and demonstrate its technical accuracy.

3.1. Scenarios used to demonstrate clinical safety and accuracy

The response rate of consultants completing the first study was 79%. The study results are in Figure 1. All scenario results (i.e. the majority consensus decision), whether to transfuse or not, were in line with BSH and NICE guidelines.

In scenarios M2, M7, S10, S14 and M13 one consultant disagreed with the others about whether to transfuse or not. However, in scenarios M2 and M13 the overall mean decision certainty was above 50% (52% and 55%, respectively) so the consensus view was accepted. In scenarios M7 and S14 there was a majority agreement on patient management greater than two thirds (as three out of

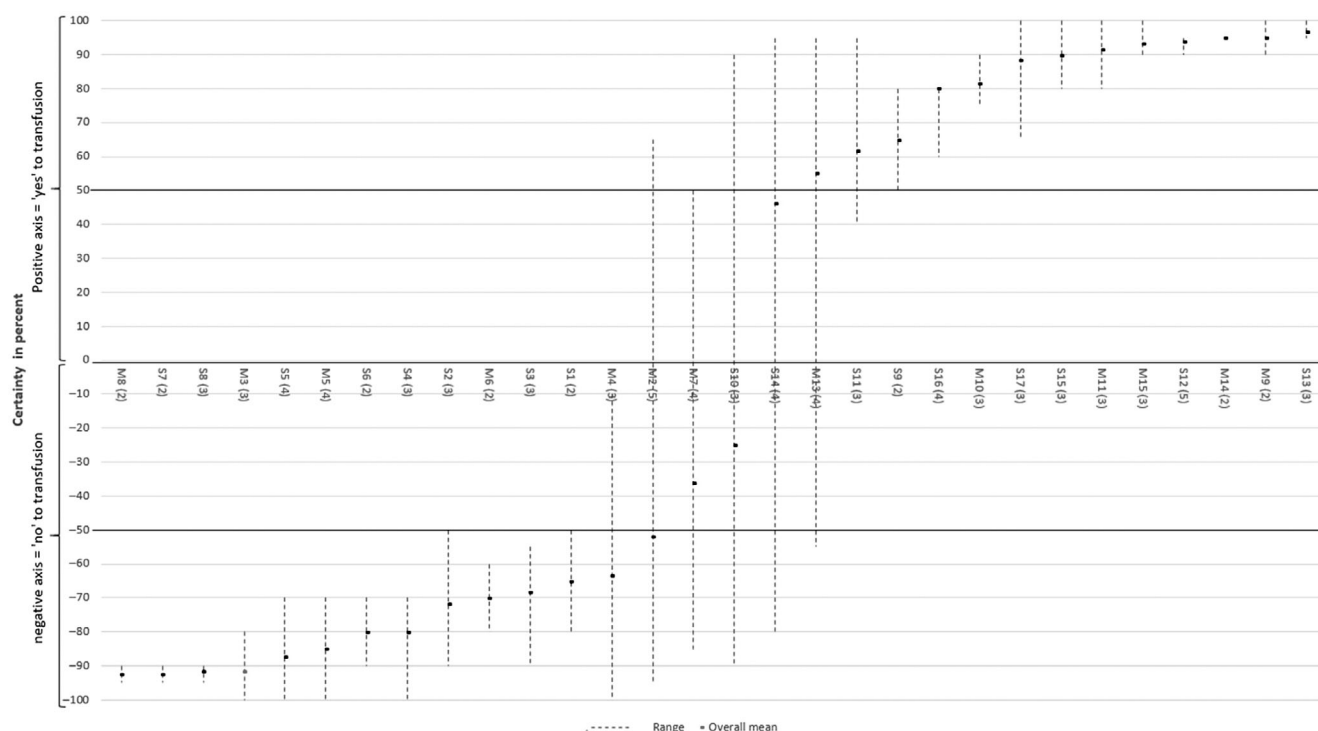


FIGURE 1 Outcome of accuracy study. This shows the mean (black dot) and range (dotted line) certainty per scenario. Scenarios were plotted across the x-axis, middle of the graph at zero, with the number of responses received (*n*). Decisions to transfuse or not to transfuse were placed on the positive or negative range on the y-axis, respectively. Scenarios were sorted left to right based on mean certainty scores from no transfusion to transfusion. Whereas some of the 30 scenarios showed good consensus, indicated via a tight range, others showed a very wide range of consultant views on whether to transfuse (above the midline) or not (below the midline). Scenario details can be seen in Appendix S1.

four consultants agreed) with the decision to transfuse or not. However, in scenario S10, only two out of three respondents agreed with management so this scenario was excluded from further analysis.

Additional comments indicated that the scenarios which showed low consensus (i.e. scenarios M2, M7, S10, S14 and M13) may have been too brief and more details on the patient condition would be needed for clarity.

In the hands of the consultants, the App was unable to provide clear guidance on 6 (21%) of the remaining 29 scenarios (M6, M13, M14, S4, S11, S14). In some of these scenarios, the consultants did not record the App's output while in others the consultants recorded four verdicts, two to transfuse and two not. These six scenarios were therefore also excluded from further analysis.

App performance was examined in the remaining 23 scenarios against the consultant's consensus decision. The App agreed with the decision of the consultant experts in 21 of 23 (91%) scenarios where it gave a clear output. For two scenarios (S2 and S5), the App recommended a transfusion but experts disagreed, giving two false-positive errors. The sensitivity of the App was therefore 100% with a specificity of 85%. This part of the study demonstrated that some App data entry screens needed improvement and/or the advice wording required modification. These changes were made and the results checked by the project group using the same scenarios before the next phase of the study.

3.2. Impact on simulated junior and middle-grade doctor decisions using scenarios

Twenty-six responses were received from junior doctors in two NHS Trusts (note: not all doctors completed all questions). Due to the differing approaches taken to recruit participating doctors, denominator data were not available to allow calculation of a response rate. An overview of participant demographics and their use of existing clinical apps can be seen in Table 1.

Among the participants, 81% of them said that they use clinical Apps, the most popular purpose being 'quick access to guidelines' with the most common in this category being the BNF App. This was followed by Apps that provide 'directory' information¹⁴ and those that support medical diagnoses and clinical decision making.

Table 2 shows the accuracy of the junior doctors on these scenarios both before and after using the App, and the impact of the App on simulated clinical decisions in these 26 participants, based on the percentage of correct decisions made (i.e. a decision compliant with guidelines).

Scenarios, S9, S1 and S8, led to a 100% correct decision making, with-App, from a high level no-App. Scenario S7 started at a lower correct decision level which did increase, but not to 100%. Three respondents, for scenario S7, who gave an incorrect decision no-App still gave an incorrect decision with-App. Participant comments showed that: the App gave them more things to consider including

**TABLE 1** Characteristics of participants

	Clinical Role						Overall
	FY1	FY2	ST4	ST5	ST6	SPR	
Number of participants	8	5	6	3	3	1	26
% Rotation in haematology (n)	0% (0)	20% (1)	100% (6)	100% (3)	100% (3)	100% (1)	54% (14)
% Who felt they had sufficient transfusion training (n)	25% (2)	40% (2)	83% (5)	100% (3)	100% (3)	100% (1)	62% (16)
% Who use clinical Apps (n)	100% (8)	100% (5)	67% (4)	33% (1)	67% (2)	100% (1)	81% (21)

TABLE 2 Comparison of participant decision adherence with RBC transfusion guidelines without (No-App) and with (With-App) access to the App

Scenario ID	Percent correct red blood cell transfusion decision compared to guidelines								
	No App			With App			Difference		
	Junior doctor	Middle-grade doctor	Overall	Junior doctor	Middle-grade doctor	Overall	Junior doctor	Middle-grade doctor	Overall
S9	92% (12/13)	80% (8/10)	87% (20/23)	100% ^a (12/12)	100% ^b (6/6)	100% (18/18)	8%	20%	13%
S1	85% (11/13)	100% (10/10)	91% (21/23)	100% ^c (12/12)	100% (6/6)	100% (18/18)	15%	0%	9%
M15	92% (12/13)	100% (10/10)	96% (22/23)	33% ^d (3/9)	75% ^e (3/4)	46% (6/13)	-59%	-25%	-50%
S8	100% (13/13)	90% (9/10)	96% (22/23)	100% (12/12)	100% ^f (6/6)	100% (18/18)	0%	10%	4%
S7	62% (8/13)	50% (5/10)	57% (13/23)	78% ^g (7/9)	75% ^h (3/4)	77% (10/13)	16%	25%	20%
Overall			85%			88%			3%
Overall Excluding Scenario M15			83%			96%			13%

Note: FY1 and FY2 doctors were grouped as 'Junior Doctors' and ST4, ST5, ST6 and SPR were grouped as 'Middle-Grade Doctors'.

^aThe participant with incorrect decision No App gave a correctly decision With App.

^bOf the two participants that gave an incorrect decision No App: one gave a correct decision With App and other did not record any further decisions With App.

^cBoth participants with an incorrect decision No App went on to record a correct decision With App.

^dThe one participant with the incorrect decision No App, did not complete the question With App, however, did leave a comment stating they did need to reconsider their decision With App. Six participants recorded a correct decision No App but recorded an incorrect decision with App.

^eOne participant who previously gave the correct answer No App, gave incorrect answer With App.

^fParticipant who recorded an incorrect decision No App, subsequently recorded the correct decision With App. Four respondents who recorded a correct decision No App, did not complete the questions With App.

^gOf the participants who did not complete the response With App: 3 previously gave the correct decision and one gave an incorrect decision. Two participants who gave an incorrect decision No App subsequently gave a correct decision With App. Two participants who gave an incorrect decision No App still gave an incorrect decision With App.

^hOf the five participants who recorded an incorrect decision No App: two did not record a decision With App (along with four who gave a correct decision), two gave a correct decision With App and one still recorded an incorrect decision With App.

the need to investigate underlying causes of anaemia; users wanted more information or lab results before making a final decision and users wanted more information for ongoing management. Their comments also stated that they would still give a transfusion whilst waiting for additional lab test results, which was against guidelines.

The mean participant decision accuracy including both junior and middle grade doctors on the five scenarios with-App was 85% and no-App was 88%, indicating that the App had no effect on the level of correct decision making. This was due to scenario M15 reducing the level of correct decisions made (from 96% no-App to 46% with-App), an example of automation bias.

Additional feedback on scenario M15 from doctors who went on to make the incorrect decision with-App, showed they wanted more information and to consider additional symptoms which affect

palliative care patients, they also noted the need for further investigations and that the transfusion decision will also depend on their local policy.

However, in the first study, scenario M15 had strong consultant haematologist consensus in line with guidelines, and strong agreement with their decision when compared to the App. Revisiting the outcome screen, the user was correctly told to transfuse but was initially presented with other options to optimise the patient and directed to follow local policies. It is possible that the mistakes observed may have been because more experienced haematologists consider greater options for intervention in addition to transfusion, which may not occur to doctors with less haematology experience. Sixty-nine percent of junior and middle-grade doctor participants indicated they did not feel they had sufficient transfusion training. From this result, it was

TABLE 3 Chi-square test for overall results excluding scenario M15

Actual results—overall			
App	Decision		Total
	Success	Failure	
With App	64	3	67
Without App	76	17	92
	140	19	159
Expected results under H_0 —overall			
App	Decision		Total
	Success	Failure	
With App	58.99	8.00	67
Without App	81.00	10.99	92
	140	19	159
Chi-square test for independence			
<i>p</i> Value	0.013		

Note: The chi-square test for independence requires that the highlighted cells in the Expected Results have a sample size of 5 or more, which was fulfilled by the data.

clear that the outcome screen of the App required to be changed to maintain and improve decision making for palliative care patients.

Therefore, in a post hoc sensitivity analysis, scenario M15 was excluded from the results of a chi-squared test (Table 3). The *p*-value of 0.013 indicates that there is evidence to reject the null hypothesis (H_0): that the App scenario and the resulting decision are independent. In other words, there is evidence to say that using the App had a potential to influence correct decision making in this sample of junior and middle grade doctors, once the scenario that gave misleading results was excluded.

4. USER SATISFACTION

During the scenario-based study, the 26 participants were asked to provide general feedback on the App. All respondents found the App easy and quick to use and found the terminology understandable. Other positive feedback included that the App had a clear layout with intuitive menus and the pop-up prompts (e.g. Transfusion Associated Circulatory Overload risk flag) were well liked.

The App was criticised for: lots of complex patient scenarios not being covered; duplication and complicated if the patient does not fit pre-defined criteria. Whereas the App only allowed the input of predefined variables that impacted the decision to transfuse red cells, the text did encourage users to look at the wider clinical situation and consult both local guidelines and local haematology/transfusion experts.

Participants were asked if they would recommend the App to colleagues. 94% of respondents (17/18) would recommend the App to other clinical colleagues. When asked if the App gave understandable advice, 94% of respondents (16/17) agreed that the advice given by

the App made sense. Two different respondents stated 'no' to these two questions but did not provide reasons for their answers.

Overall, most (17, 90%) of 19 end-users stated that they would use the App in clinical practice, with 6 (32%) stating that they would sometimes use it. Of the two (11%) who said they would not use it, one (ST4 grade) stated the App was more 'suitable for junior team members and non-haematologist' and the other (ST5 grade) stated 'many patients will fall into the "medical other" section where the advice is "investigate cause" and does not tell you whether or not to transfuse'. The ST5 went on to give positive comments on the App's build and usability but also stated 'surgical sections seemed useful and more precise'.

5. DISCUSSION

The successful co-development of the Blood Choices App involved a multi-disciplinary team of clinical, digital health and evaluation experts, end-users and information technology specialists. At the time, other blood transfusion Apps providing calculators and generic access to guideline information were becoming available. However, this was the first known attempt to provide safe and appropriate smartphone-based transfusion decision support by directing users to the relevant part of a guideline based on patient variables.

The Blood Choice App's performance was examined in 23 varying clinical scenarios against the consultants' consensus decision, defined as a mean decision certainty above 50% or a greater than two thirds majority. This assessment showed that the App agreed 91% of the time with the decision of experts in the field with a sensitivity of 100% and specificity of 85%.

The haematology consultants selected for this study were blood transfusion leads and would have been well versed in, or authors of, national RBC guidelines. The less than 100% consensus highlights that making a decision to transfuse RBCs is often complex, and this complexity has been demonstrated during the development of behavioural interventions to change transfusion practice.¹⁵ Other factors, such as the clinicians' individual risk versus benefit tolerance, could have influenced their decisions on scenarios.

This may also reflect the fact that the decision to transfuse depends upon individual patient factors rather than a set of pre-specified triggers, which has been highlighted in previous work on decision support tools for transfusion.¹⁶ It is also possible that the way in which the scenarios were written or presented were not understood as intended.

Consultant haematologists with responsibility for transfusion are linked with the wider blood transfusion community to keep up-to-date with current evidence base, and tend to influence the transfusion culture within their healthcare organisation, e.g. they are the final arbiter for decisions by more junior or non-specialised doctors. Past audits of clinical practice have shown variation in transfusion decisions, which was also observed in this study. Further research should examine the drivers of individual consultants to follow and implement guidelines and effective strategies for guideline implementation within different clinical areas.

For the study of App impact on simulated junior doctor decisions, only five scenarios were used due to expected engagement difficulties with junior doctors and their time limitations for additional non-vocational activities. Had more time and a larger population of junior doctors been available, additional scenarios could have been tested.

Participant responses to scenario M15 demonstrated “automation bias”,¹⁷ where users followed the advice of the App even when its advice was mistaken. As expected, the change in decision was worse in junior doctors (from 92% to 33%), possibly due to their inexperience compared to middle grade doctors (100%–75%).

As scenario M15 highlighted the need to make changes to the App, the App was not made available after the pilot. This result highlighted the importance of clinical App evaluations. The information on the outcome screen for this scenario was viewed, updated and then discussed to remove ambiguous advice. However, due to resource limitations, the inconclusive result was not retested.

Scenario S7 demonstrated that junior doctors will use the information in the App as well as their own clinical judgement to make an informed decision whether to transfuse. Even if these decisions are outside guidelines, doctors made the decision after viewing guidelines so it could be argued that they are better informed. Ideally guidelines attempt to make clinicians aware of recommended evidence-based advice. However, as there will always be patients who fall outside the scope of evidence-based guidelines, being informed of the recommendations but still using clinical judgement is how the best individual care can be delivered. Technology should support doctors in their clinical decisions, and not attempt to replace them.

Patient Blood Management initiatives have been more successful within surgery than within medical settings.¹⁸ During this study, slightly more surgical scenarios were presented to the consultant haematologist group, however, three of the five scenarios with low consensus were medical. In addition, the one scenario associated with malfunction from the junior doctor study was also medical. Finally, user feedback indicated that more factors and information were needed when making a decision in medical scenarios and the App was more suited for surgical scenarios. These results suggest that more research is needed into how best to factor in, and the rational for, the wider considerations needed for making decisions around medical RBC use.

Overall, this evaluation showed that the Blood Choices App does have the potential to improve adherence to RBC guidelines, with further iterations. Eighty-one percent of potential end-users surveyed said that they use clinical Apps. If end-users are satisfied with the App, they may be more likely to use it. With 58%⁸ of end-users saying they would use the App in clinical practice, and 32% saying they will use it sometimes, Apps, such as the Blood Choices, could provide an innovative solution to influence the rate of RBC transfusions.

5.1. Strengths of the study

The methods used for undertaking the accuracy and simulated impact studies were in line with published literature.^{12,13}

A large number of scenarios covering a range of transfusion settings were developed based on areas of interest from past clinical audits and covered by national guidelines. To remove any bias from the project group, the accuracy evaluation tested the App's decisions against the outcomes referenced by the consensus decision of haematology consultants who are leads for transfusion. Evaluating accuracy in this way should form a large part of App safety assessment. In addition, this pilot went further than most published App studies to explore the potential impact of the App's output on decision making by the targeted users, using a smaller number of scenarios. This uncovered the risk of automation bias in one scenario, which resulted in the App not being made available for any wider launch or evaluation following this pilot work. This finding clearly demonstrated the importance of clinical App testing with intended users.

5.2. Study weaknesses

The study was designed on the assumption that practice variations were due to the lack of awareness of evidence-based guidelines, which the App can support, and not because clinicians may not trust guidelines or choose not to follow them. Future research on guideline awareness and reasons for rejecting their advice is, therefore, recommended.

It should also be noted that the sample size for simulated decision-making using scenarios was limited both by the number of scenarios used and the number of participants. Asking participants what they would do related to transfusion was not a mandatory question, leading to a drop in participant response rates in later questions, as some participants skipped the answer with-App. Responses from haematology consultants were also not mandatory, leading to the exclusion of 6 scenarios to compare the Apps responses, which may have impacted the Apps performance.

The App's design was targeted towards junior and middle-grade doctors, regardless of their training in haematology. The difference in understanding of RBC guidelines post haematology training could be a consideration for future studies and App design.

Understanding the reasons for junior doctor non-adherence to guidelines could provide further insights to improve the efficacy of a future App. In addition, increasing the expert haematology consultant pool to increase the number of responses per scenario would have increased majority responses and ensured that all scenarios could have been included in the analysis.

This study used up to five haematology consultant decisions on scenarios as the consensus decision, as it is likely they would have been asked by more junior medics how to manage the patient. Given the variable number of consultants responding to each scenario and our limited resources, we decided to take an overall mean decision certainty above 50% as the reference standard without carrying out a detailed measurement study in advance, to ensure that this was a reliable, valid measure.¹⁹ Future studies should explore how to better define consensus consultant decisions using psychometric techniques to explore the impact of changing the mean decision certainty

threshold for exclusion and the number of consultant responses required to define a reliable, valid gold standard.

Even though the level of potential user satisfaction was high, we only tested the impact of the App with case scenarios on simulated decisions, not on real patient decisions. Future satisfaction scores should also be evaluated from App store user ratings.

The study of junior doctor responses to the App may have been subject to a Hawthorne effect or social response bias. It is difficult to test the App on all patient cases requiring a transfusion decision to be made. The only valid way to determine if the App can improve transfusion decisions in real patients would be to carry out the study as a randomised controlled trial within a clinical setting. This should be accompanied by a process evaluation using qualitative discussions with doctors who have just made a decision for RBC transfusion and ask them to review that decision with the App and then ask if they would do anything differently. Google Analytics could also be used in future evaluations to understand how users navigate the App to get to the outcome screens, to guide further development.

It is recommended that healthcare organisations, professional bodies and individual clinicians push for evaluation of clinical Apps to demonstrate their safety and accuracy before they are promoted for use. This evaluation is important because, as demonstrated in this study when advice was not clear, apps can lead to incorrect decisions.

Since taking the App off the market following the simulated impact study, the App was not re-issued for download, mainly due to resource constraints post pilot. However, the NHSBT's Patient Blood Management Team has since worked with clinicians and invested in other Apps, such as *Blood Components*.²⁰ In addition, there were imminent changes to the MHRA medical device regulations which need to be considered before re-issue of a CE-marked App. To minimise resource allocation, the Blood Choices App is being reviewed and will be made available for use as an educational, rather than a decision support, tool.

The development of Apps can require external developers, leading to an economic challenge.

The total cost of this project was circa £47 000. Although no economic evaluation was completed, a review of literature indicates that Apps can provide a financial return as well as healthcare improvement. Bricker et al.²¹ reported accessing guidelines from Apps could provide a low-cost intervention with strong population impact. Ryan et al.²² reported whereas Apps were an additional cost for the NHS, with widespread adoption overall costs could be reduced. Liu et al.²³ reported that guideline adherence using the App was comparable to using paper plans but concluded that Apps were a cost-effective approach to self-management, leading to a reduction in standard care costs. If these findings are transferable, the successful development of RBC transfusion guidelines into an App has a strong economic argument and future health economic work is recommended.

6. CONCLUSIONS

The project delivered a *Blood Choices* App which had high end-user satisfaction. It was evaluated by experts as being accurate and safe but

due to low sample size and the use of case scenarios to study simulated rather than real transfusion decisions, it is too early to conclude if Apps, such as the *Blood Choices*, will improve RBC guideline adherence.

However, with further development and testing, Apps such as the *Blood Choices* have the potential to improve doctor's decision making. Coupled with smartphones being the most accessible form of technology globally and reports of healthcare staff being empowered to challenge inappropriate care, accessing RBC guidelines in App format is an obvious route for further development.

Critically, this study further demonstrates the need to ensure clinical Apps are robustly evaluated and that the limitations of algorithm-based Apps are acknowledged and well understood by the users and reflected through co-design development and testing.

AUTHOR CONTRIBUTIONS

Amanpreet Dhese led the project including co-authoring funding application, App development, delivery and evaluation. He drafted the manuscript, incorporated changes and submitted final. Jeremy Wyatt made a significant contribution to the design of both studies and to the data analysis. He contributed to the manuscript and edited and approved the final version. Lise Estcourt developed the App content, supported study design and developed the scenarios used. She contributed to the manuscript and read and approved the final version. Wendy McSporran developed the App content, supported study design and developed the scenarios used. She contributed to the manuscript and read and approved the final version. Shubha Allard was the overall medical lead for the project and co-authored the funding application. Contributed and gave approval to App development, validation scripts, design evaluation studies and scenarios used. She contributed to the manuscript and read and approved the final version.

ACKNOWLEDGEMENTS

The authors thank all clinical colleagues who inputted, provided and reviewed content and all those who gave up their time as study participants. Sue Katic (Senior Administrator), for organising and managing stakeholders and reporting for the entire project. Duncan Brian (Specialist Registrar in Haematology) and Gavin Cho (Consultant Haematologist), for their input into App development (content, decision logic and user journeys) and development of validation protocols. Catherine Booth (Consultant Haematologist), Nidhi Sharma (Transfusion Practitioner) and Jo McCullagh (Transfusion Practitioner), for their input and active engagement of end users to complete the study evaluations. Elisa Allen (Principle Statistician), Statistics and Clinical Studies, NHS Blood and Transplant for her support in statistical analysis. Brian Hockley (Data and Audit Manager), NHS Blood and Transplant for designing survey on *SnapSurvey*© and processing results. The NHSBT Research and Development Trust Fund and NHSBT Patient Blood Management Team for funding this project. HMA Digital for their support and engagement in developing the Blood Choices App and working with us on this project.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

ORCID

Lise J. Estcourt  <https://orcid.org/0000-0003-4309-9162>

Wendy McSporran  <https://orcid.org/0000-0002-1558-9271>

REFERENCES

1. Flodgren G, Rojas-Reyes MX, Cole N, Foxcroft DR. Effectiveness of organisational infrastructures to promote evidence-based nursing practice (review). *Cochrane Database Syst Rev*. 2012;2:1-4.
2. NCABT. Audit of blood transfusion in adult cardiac surgery – part one. *National Comparative Audit of Blood Transfusion [Online]*. 2011;1-40. http://hospital.blood.co.uk/media/26859/nca-2011_use_of_blood_in_adult_cardiac_surgery_report.pdf
3. NCABT. Audit of use of blood in adult medical patients - part two. *National Comparative Audit of Blood Transfusion [Online]*. 2013;1-43. http://hospital.blood.co.uk/media/26861/nca-2011_medical_use_audit_part_2_report.pdf
4. NCABT. Repeat Audit of Red Cell and Platelet Transfusion in Adult Haematology Patients *National Comparative Audit of Blood Transfusion [Online]*. 2017;1-83. https://nhsbt.dbe.blob.core.windows.net/umbraco-assets-corp/14912/report_26007737.pdf
5. Charani E, Castro-Sánchez E, Moore LSP, Holmes A. Do smartphone applications in healthcare require a governance and legal framework? It depends on the application! *BMC Med*. 2014;12:29.
6. Mobasheri MH, King D, Johnston M, Gautama S, Purkayastha S, Darzi A. The ownership and clinical use of smartphones by doctors and nurses in the UK: a multicentre survey study. *BMJ Innovations*. 2015;1:174-181.
7. Payne KFB, Wharrad H, Watts K. Smartphone and medical related app use among medical students and junior doctors in the United Kingdom (UK): a regional survey. *BMC Med Inform Decis Mak*. 2012;12:1-11.
8. Charani E, Kyrtatsis Y, Lawson W, et al. An analysis of the development and implementation of a smartphone application for the delivery of antimicrobial prescribing policy: lessons learnt. *Journal of Antimicrob. Chemotherapy*. 2013;68(4):960-967.
9. Panesar P, Jones A, Aidous A, et al. Attitudes and behaviours to antimicrobial prescribing following introduction of a smartphone app. *PLoS ONE*. 2016;11(4):e0154202.
10. Song Y, Oh J, Chee Y, Cho Y, Lee S, Lim TH. Effectiveness of chest compression feedback during cardiopulmonary resuscitation in lateral tilted and semirecumbent positions: a randomised controlled simulation study. *Journal of Anaesthesia*. 2015;70:1235-1241.
11. Low D, Clark N, Soar J, et al. A randomised control trial to determine if use of the iResus© application on a smart phone improves the performance of an advanced life support provider in a simulated medical emergency. *Journal of the Association of Anaesthetists of Great Britain and Ireland*. 2011;66:255-262.
12. RCP. Using apps: clinical practice guidance. Royal College of Physicians website [online]. 2015. Accessed August 11, 2018. <https://www.rcplondon.ac.uk/guidelines-policy/using-apps-clinical-practice-guidance>
13. Wyatt JC. How can clinicians, specialty societies and others evaluate and improve the quality of apps for patient use? *BMC Med*. 2018;16:225.
14. Wyatt JC, Sullivan F. 2005. Communication and navigation around the healthcare system. *BMJ*. 2005;331(7528):1325-1327. doi:10.1136/bmj.331.7528.1325
15. Francis JJ, Stockton C, Eccles MP, et al. Evidence-based selection of theories for designing behaviour change interventions: using methods based on theoretical construct domains to understand clinicians' blood transfusion behaviour. *Br J Health Psychol*. 2009;14(4):625-646.
16. Goodnough LT, Shah N. Is there a "magic" hemoglobin number? Clinical decision support promoting restrictive blood transfusion practices. *Am J Hematol*. 2015;90(10):927-933.
17. Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. *J Am Med Inform Assoc*. 2012;19(1):121-127. doi:10.1136/amiajnl-2011-000089
18. NICE Guideline, NG24. *Blood Transfusion*. NICE; 2015.
19. Scott PJ, Brown AW, Adediji T, et al. A review of measurement practice in studies of clinical decision support systems 1998-2017. *J Am Med Inform Assoc*. 2019;26(10):1120-1128. doi:10.1093/jamia/ocz035
20. Watson P, Watson D, Dhese A, et al. Improving blood-prescribing decisions: evaluating the efficacy of a blood guidelines app. *Transfus Med*. 2020;30:6-491.
21. Bricker JB, Mull KE, Kientz JA, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. *Drug Alcohol Depend*. 2014;143:87-94.
22. Ryan D, Price D, Musgrave SD, et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. *BMJ*. 2012;344(1756):1-15.
23. Liu WT, Huang CD, Wang CH, Lee KY, Lin SM, Kuo HP. A mobile telephone-based interactive self-care system improves asthma control. *Eur Respir J*. 2011;37:310-317.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Dhese AS, Wyatt JC, Estcourt LJ, McSporran W, Allard S. Insights from developing and evaluating the NHS blood choices transfusion app to support junior and middle-grade doctor decision making against guidelines. *Transfusion Medicine*. 2022;1-9. doi:10.1111/tme.12872